TESTING THE UNIVERSALITY OF THE FUNDAMENTAL METALLICITY RELATION AT HIGH REDSHIFT USING LOW-MASS GRAVITATIONALLY LENSED GALAXIES

Sirio Belli[1], Tucker Jones[2], Richard S. Ellis[1], Johan Richard[3]
Submitted to ApJ

ABSTRACT

We present rest-frame optical spectra for a sample of 9 low-mass star-forming galaxies in the redshift range $1.5 < z < 3$ which are gravitationally lensed by foreground clusters. We used Triplespec, an echelle spectrograph at the Palomar 200-inch telescope that is very effective for this purpose as it samples the entire near-infrared spectrum simultaneously. By measuring the flux of nebular emission lines we derive gas phase metallicities and star formation rates, and by fitting the optical to infrared spectral energy distributions we obtain stellar masses. Taking advantage of the high magnification due to strong lensing we are able to probe the physical properties of galaxies with stellar masses in the range $7.8 < \log M/M_\odot < 9.4$ whose star formation rates are similar to those of typical star-forming galaxies in the local universe. We compare our results with the locally determined relation between stellar mass, gas metallicity and star formation rate. Our data are in excellent agreement with this relation, with an average offset $\langle \Delta \log(\mathrm{O/H}) \rangle = 0.02 \pm 0.08$, suggesting a universal relationship. Remarkably, the scatter around the fundamental metallicity relation is only 0.24 dex, smaller than that observed locally at the same stellar masses, which may provide an important additional constraint for galaxy evolution models.

Subject headings: galaxies: abundances – galaxies: evolution – galaxies: high-redshift – gravitational lensing: strong

1. INTRODUCTION

The gas-phase metallicity represents a fundamental property of galaxies and can be used to investigate the complex physical processes that govern galaxy evolution. It mainly traces the star formation history, as metals produced in stars are ejected into the interstellar medium (ISM), but the exchange of material between the galaxy and the intergalactic medium (IGM) also plays an important role. The accretion of metal-poor gas from the IGM can dilute the metal content of the gas in a galaxy. Also, stellar winds can substantially lower the metallicity by ejecting metals.

Despite the complexity of these processes, a clear relation between galaxy luminosity and metallicity has been known since the work of Lequeux et al. (1979). Recently, thanks to the vast amount of spectroscopic and photometric data available from the Sloan Digital Sky Survey (SDSS), it has become clear that the physical parameter that correlates most strongly with metallicity is the galaxy stellar mass (Tremonti et al. 2004). This mass-metallicity relation, in which galaxies of higher masses contain larger metallicities, is remarkably tight over 3 orders of magnitude in stellar mass, with a dispersion of only 0.10 dex in metallicity.

A natural explanation for the observed mass-metallicity relation is the outflow of metal-enriched gas driven by star formation. Because of the lower gravitational potential, low-mass galaxies lose a higher frac-tion of their gas, with a consequent decrease in metallicity (Larson 1974; Garnett 2002; Tremonti et al. 2004). An alternative possibility is that lower mass galaxies are less metal-rich because their star formation history has been developed more gradually (Ellison et al. 2008), in agreement with the now-familiar effect of *downsizing* (Cowie et al. 1996).

Different models of galaxy formation and evolution are able to match the mass-metallicity relation in the local universe, but have dissimilar predictions for high redshift galaxies (e.g. De Lucia et al. 2004; Davé & Oppenheimer 2007; Tassis et al. 2008; Davé et al. 2011a; Yates et al. 2012). Observing the redshift evolution of the mass-metallicity relation can therefore differentiate these models. Observations at different redshifts have shown a clear evolution with cosmic time, with lower metallicity at higher redshift, for a fixed mass (Savaglio et al. 2005; Erb et al. 2006a; Maiolino et al. 2008; Mannucci et al. 2009; Zahid et al. 2011; Yuan et al. 2013).

However, it is important to recognize that high-redshift studies target galaxy populations that are different from those found typically in the local Universe. The evolution of the mass-metallicity relation could then be the result of a selection effect rather than a change in the physical properties of the galaxies with cosmic time.

Among the differences between local and high-redshift galaxies, star formation activity is one of the most important. At earlier cosmic times, the star formation rate (SFR) was on average much higher than today, because galaxies contained a larger amount of cold gas. Also, most of the high-redshift surveys are magnitude-limited in the rest-frame UV, and therefore tend to select galaxies with high SFR. The combination of these two effects makes it very difficult to compare the metallicity of galaxies at different redshifts with the same stellar mass

[1] Department of Astronomy, California Institute of Technology, MS 249-17, Pasadena, CA 91125, USA
[2] Department of Physics, University of California, Santa Barbara, CA 93106, USA
[3] Centre de Recherche Astrophysique de Lyon, Université Lyon 1, 9 Avenue Charles André, 69561 Saint Genis Laval Cedex, France

and star formation rate. It is then essential to study the relation between SFR and metallicity, since this could have important consequences on the interpretation of the observed evolution of the mass-metallicity relation.

In fact, Mannucci et al. (2010) found that the local mass-metallicity relation is different for samples of galaxies with different star formation rates. Furthermore, they showed that the SDSS galaxies lie on a tight 3D surface in the mass-metallicity-SFR space, with a dispersion of only 0.053 dex in metallicity (see also Lara-López et al. 2010). According to this *fundamental metallicity relation* (FMR), at fixed stellar mass SFR and metallicity are anti-correlated. If this relation holds independently of cosmic time, a galaxy population at high redshift will tend to have a low average metallicity because of its high SFR. In this scenario the evolution of the mass-metallicity relation is driven by the shifting of galaxy populations on the SFR-mass plane, rather than being directly caused by the evolution of some physical process. Clearly, testing the fundamental metallicity relation at different redshifts is of primary importance.

The redshift evolution of the FMR was first explored by Mannucci et al. (2010) using samples from the literature, and they concluded that the local relation is a good fit for any star-forming galaxy up to $z \sim 2.2$. But high-redshift observations are biased towards high-SFR galaxies, and a direct test using galaxies with the same range of star formation rates that is seen in the SDSS sample (SFR < $10 \ M_{\odot}/\mathrm{yr}$) is still lacking. Additional difficulties come from the fact that to measure the metallicity one needs the rest-frame optical emission lines, which at $z \sim 2$ are redshifted into the near-infrared, a spectral region where sky emission is strong.

One way to probe lower star formation rates with the current technology is to take advantage of strong gravitational lensing. The magnification induced by foreground galaxy clusters allows one to reach faint objects, corresponding to stellar masses and SFRs (on average) lower than the values achievable without lensing. Recent studies of the FMR for high-redshift lensed galaxies found a general agreement with the local relation, although with a very large scatter (Richard et al. 2011; Wuyts et al. 2012b; Christensen et al. 2012).

Testing the universality of the fundamental metallicity relation is important not only for understanding the evolution of the mass-metallicity relation, but also to constrain models of galaxy evolution. For example, Davé et al. (2011a) consider a simple model in which inflow, outflow and star formation are in equilibrium and determine the gas metallicity. This scenario can qualitatively explain the dependence of the mass-metallicity relation on the star formation rate: at a fixed stellar mass a high SFR is caused by a large inflow, that in turn implies a low metallicity. If a galaxy is perturbed, e.g. by a merger, it will move away from the FMR and, after some time, will return to the equilibrium configuration. This equilibrium timescale determines the scatter in the mass-metallicity-SFR relation. So long as the yield of metals per unit star formation and the mass loading factor (i.e. the ratio between outflow and star formation) are constant, the equilibrium relation implies a universal FMR independent of redshift. Although this and other simple analytic models (Dayal et al. 2012; Davé et al. 2012) succeed in explaining the local fundamental metallicity

relation, we are still far from a detailed understanding of the relevant physical processes. High-redshift observations are essential for quantitative tests of hydrodynamical simulations of galaxy evolution.

In this work we study a sample of low-luminosity, $z \sim 2$ gravitational arcs with magnification factors of ~ 10–100. We used the Triplespec spectrograph on the Palomar 200-inch telescope that features a good sensitivity and covers the full near-infrared wavelength range. This characteristic makes it an ideal instrument for such a study, because it allows us to observe all diagnostic lines of interest simultaneously. In addition to the efficiency of observation, a particular benefit is the ability to measure emission line ratios, the diagnostics of gas-phase metallicities, in a single exposure, mitigating uncertainties that arise from variations in weather conditions. Also, we mainly rely on the emission lines [OIII]$\lambda 5007$ and Hβ for measuring the metallicity, and not on the [NII]$\lambda 6584$/Hα ratio, so that we are not limited by the requirement of detecting the faint [NII]$\lambda 6584$ line. For these reasons we probe stellar masses and star formation rates that are on average lower than the ones of previously studied samples of lensed galaxies.

The sample of gravitational arcs, the spectroscopic observations, and data reduction are described in Section 2. In Section 3 we present the photometric measurements and the fitting of the spectral energy distribution, while in Section 4 we explore the galaxy physical properties using the measured line fluxes. We discuss the constraints of these measurements on the evolution of the fundamental metallicity relation in Section 5, and the implications for galaxy evolution models in Section 6. We assume a Λ cold dark matter (ΛCDM) cosmology with $\Omega_{\Lambda} = 0.7$, $\Omega_{\mathrm{m}} = 0.3$ and $H_0 = 70 \ \mathrm{km \ s^{-1} Mpc^{-1}}$. Magnitudes are given in the AB system.

2. DATA
2.1. *Sample Selection*

We selected our sample of gravitational arcs from the literature according to the following three criteria.

First, we considered only galaxy clusters with a well-constrained lens model. This allowed us to select arcs of known magnification $\mu \gtrsim 10$.

Second, the observed (i.e. not corrected for lensing) arc magnitude must be $R \lesssim 23$, so that observations with Triplespec at Palomar are feasible. This means that we can probe intrinsic magnitudes $R \gtrsim 25.5$, fainter than the limits of typical non-lensing surveys.

Third, the arc should have a known spectroscopic redshift such that the emission lines from [OII] to Hα fall in the wavelength range observable with Triplespec. The ideal range is $2 < z < 2.5$, but we can measure metallicities and star formation rates using lines available for galaxies from $z \sim 1.5$ to $z \sim 3$.

This selection provides us with 10 sources viewed through 10 distinct clusters (see Table 1). For the arc in RXJ1720 only Hα is observable because the other diagnostic lines are unfortunately obscured by night sky emission. We do not attempt any analysis on this object, but we include it in Table 1 because ours is the first redshift measurement for this arc obtained from a rest-frame optical emission line.

From now on we will refer to the gravitational arcs

TABLE 1
Sample of Gravitational Arcs

Cluster	Arc	μ^a	Reference	Run	Exp. Time	Redshift[b]
A611	1.2, 1.3	19.6 ± 3.0	Newman et al. (2009)[c]	B	1h20min	1.4902
RXJ2129	1.1, 1.2	61 ± 17	Richard et al. (2010)	A	5h30min	1.5221
A1413	2.1, 2.2	23.9 ± 6.4	Richard et al. (2010)	B, D	3h00min	2.0376
A1835	7.1	88 ± 30	Richard et al. (2010)	B	1h45min	2.0733
RXJ1720	1.1, 1.2	22.5 ± 9.1	Richard et al. (2010)	A	4h00min	2.2200
A773	1.1, 1.2	27.9 ± 7.9	Richard et al. (2010)	C	3h15min	2.3032
MACS0717	13.1	7.2 ± 3.0	Limousin et al. (2011)	C, D	5h30min	2.5515
A383	3C, 4C	23.9 ± 3.3	Newman et al. (2011)	C	2h00min	2.5771
A1689	1.1, 1.2	57 ± 23	Coe et al. (2010)	D	6h30min	3.0421
A1703	3.1, 3.2	46 ± 20	Richard et al. (2009)	B	3h30min	3.2847

[a] Gravitational magnification. For multiply imaged sources, this is the sum of the magnifications.
[b] Redshifts measured from [OIII]λ5007, except for RXJ1720, for which Hα was used. The uncertainties are always less than 0.0003.
[c] Magnification factor calculated after updating the lensing map with the new arc redshift, see Section 4.

TABLE 2
Observing runs

Run	Date	Seeing (arcsec)
A	2010 August 23, 25, 26	0.9 - 1.3
B	2011 April 10, 11, 12, 13	0.9 - 1.3
C	2012 January 12, 13, 14	1.1 - 2.0
D	2012 April 29, 30, May 1	0.9 - 1.5

by the names of the corresponding galaxy clusters, e.g. A1835 for A1835 arc 7.1.

2.2. Spectroscopy

All spectroscopic data were taken with Triplespec on the 200-inch Hale Telescope at Palomar Observatory over the course of four observing runs (see Table 2). Triplespec is a near-infrared cross-dispersed spectrograph that simultaneously covers the wavelength range 1–2.4 μm with a resolution R $\sim$ 2700 (Herter et al. 2008).

The 1×30 arcsec long slit was positioned on the targets as shown in Fig. 1 via a blind offset from a bright star. For each target we typically undertook many exposures of 300–450 seconds each, using a two-point dithering pattern. The position of the target along the slit and the dithering offset were carefully chosen for each arc, avoiding any overlap of the arc with foreground cluster galaxies, and leaving enough blank sky along the slit to reliably measure and subtract background emission. When it was possible to arrange a multiply imaged system in the slit, the spectra, if resolved, were reduced and extracted separately and then combined.

The spectroscopic data were reduced using a modified version of *Spextool* (Cushing et al. 2004; Vacca et al. 2004). For each target we extracted the spectrum from each A−B pair and then combined the 1D spectra. The extraction aperture was defined using the [OIII] emission profile in the stack of many A−B pairs, since the line emission is rarely detected in single frames.

Flux calibration and correction for telluric absorption were performed using Elias et al. (1982) A-type standard stars. Although absolute flux calibration can be very uncertain, it affects only the emission line fluxes but not their ratios, which are used in calculating gas-phase metallicities. This is one of the main advantages of using Triplespec, which allows one to observe the entire near-infrared spectrum at once. However, SFR measurements are affected by absolute flux calibration, that therefore needs to be carefully quantified.

The accuracy of the absolute flux calibration is limited by many factors. First, the observing conditions were not always photometric, and a good fraction of our observations were affected by the presence of clouds. This problem is mitigated by the fact that each target was observed on more than one night, and we scaled the different frames to match the line fluxes observed during the best conditions. The second issue arises from the fact that every few minutes the pointing was changed according to the dithering pattern. As a result, the slit alignment is not identical in each frame, and the observed flux changes as a function of the centering of the target. We tested the significance of this effect using the standard star observations, where the target is bright enough to compare the flux in different frames. The discrepancy in the absolute flux between different frames is typically less than 50%. This represents an upper limit on the flux uncertainty, since this random effect is attenuated by averaging together multiple frames for each standard star. Also, the science observations were made on much longer timescales, and the guiding was overall very stable, as we could check from the guider images taken during the exposure. Another possible source of uncertainty is the highly variable seeing. However, the difference in seeing (which was almost always larger than the slit width) between the science target and the standard star observations has a much smaller effect than the slit misalignment. Since it is difficult to quantify the real uncertainty, we conservatively use the value of 50% as fractional error in the flux calibration. This corresponds to a contribution of 0.22 dex to the uncertainty in the SFR.

The calibrated spectra together with their 1-σ error are shown in Fig. 1. The emission line [OIII]λ5007 is well-detected in each spectrum; other observed emission lines are Hα, Hβ, [OIII]λ4959, and [OII]λ3727, 3729. The fainter lines [NII]λ6584 and [NeIII]λ3870 are detected only in a few cases. The continuum emission from the arcs is never detected, but the residuals from sky subtraction or the emission from foreground galaxies can cause the observed continuum to be different from zero.
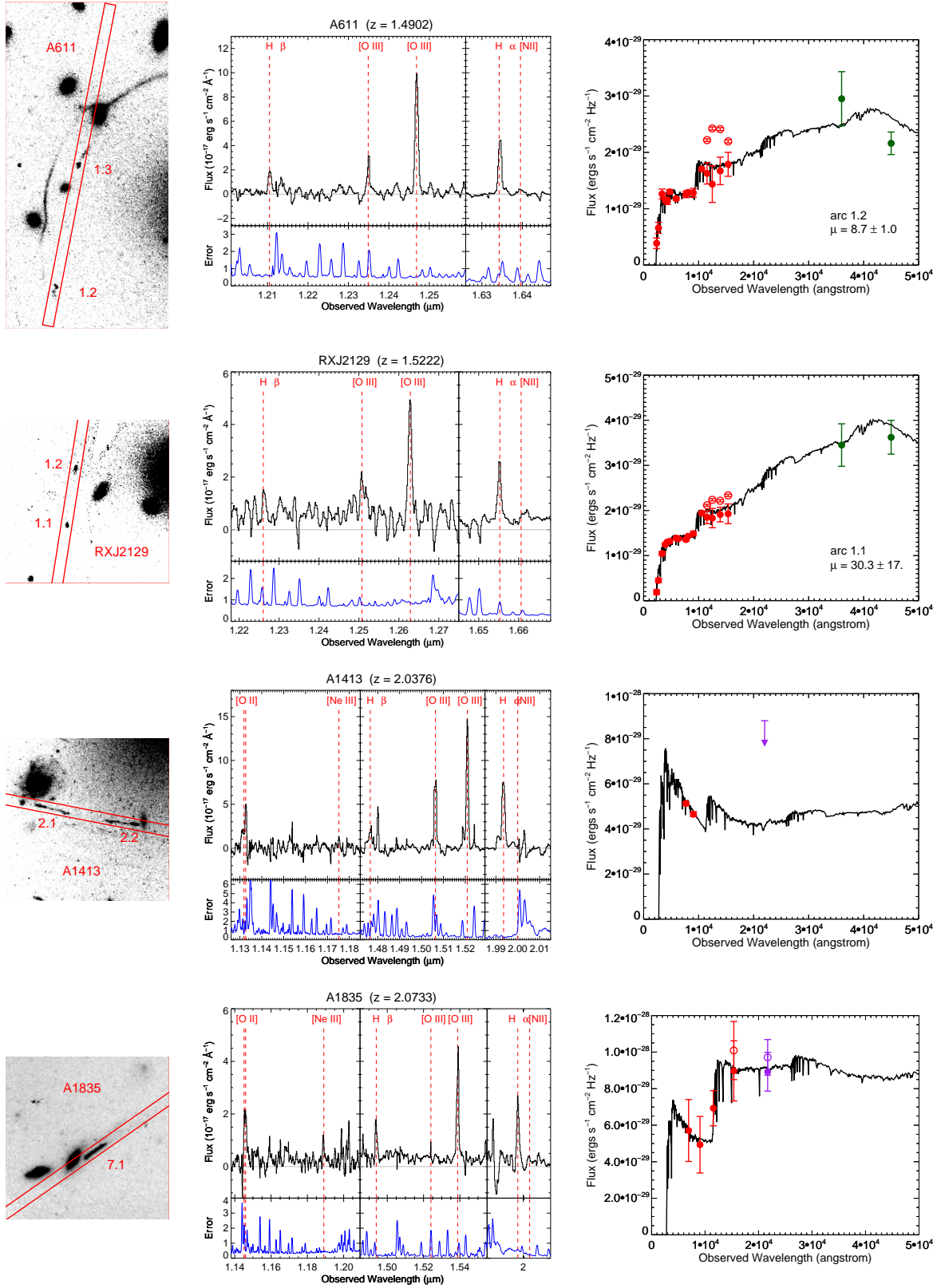
Belli et al.

FIG. 1.— *HST* image stamp, Triplespec spectrum, and SED fit for each gravitational arc. *Left:* The *Hubble Space Telescope* images (F702W or F775W) show the position of the Triplespec slit. Each square is 15 arcseconds on the side. *Center:* Each spectrum has been inverse-variance smoothed using a 5-pixel window, and the bottom panels show the 1-σ error in the same units as the flux. When multiple images are present on the slit, the spectrum shown is the combined spectrum. *Right:* The photometry from (observed) UV to infrared is plotted as filled points, and the color corresponds to the type of data: red for *HST*, purple for ground-based near-infrared, and green for *Spitzer* IRAC images. The empty points are photometric measurements not corrected for emission line flux (see Section 3.2), and the best-fit synthetic spectrum is shown as a solid line. For objects with two gravitational images, the SED is plotted only for the one that is less contaminated by foreground galaxies, and its magnification factor is shown. Spectra and photometry are not corrected for the lensing magnification.
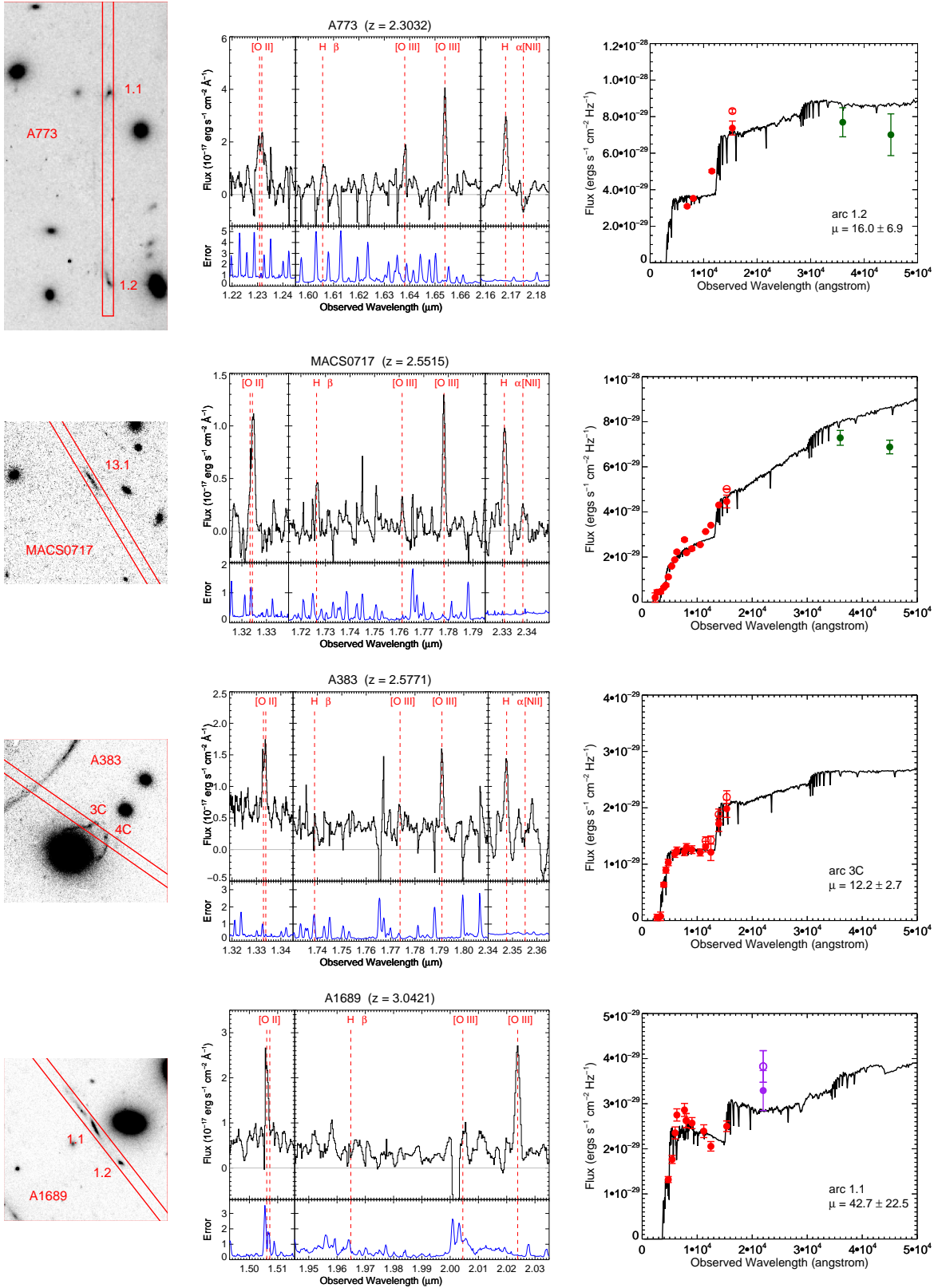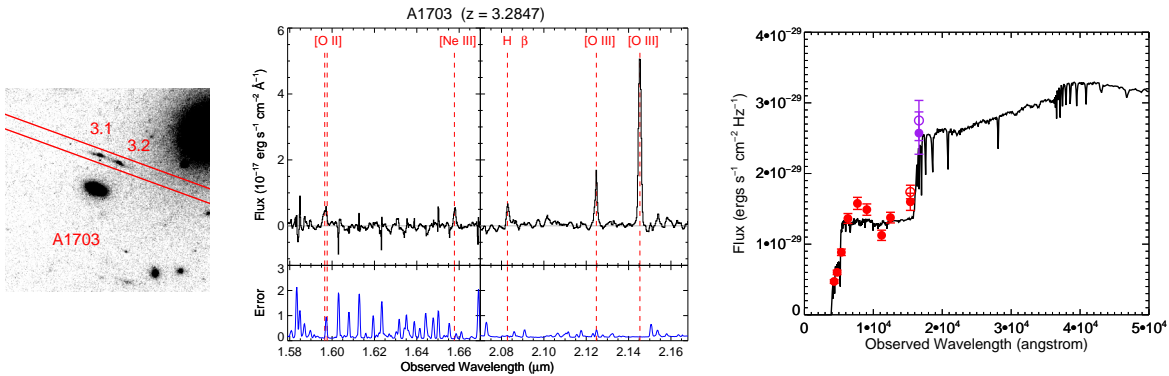
Fig. 1.— Continued.

FIG. 1.— Continued.

2.3. *Imaging*

We now discuss the imaging data for our sources which will provide the essential ingredients for measuring the stellar masses and other physical properties. To accurately derive the stellar mass it is necessary to sample the spectral energy distribution (SED) redward of the rest-frame Balmer break. For $z < 2$ objects optical imaging is sufficient, but at higher redshift infrared data are needed.

2.3.1. *Archival Data*

Since we chose well-studied galaxy clusters, space-based images from *Hubble Space Telescope* (*HST*) and *Spitzer* observations are available for each target in our sample.

For four of the objects (A383, MACS0717, RXJ2129, and A611) we used publicly available data from the Cluster Lensing and Supernova Survey with Hubble (CLASH, Postman et al. 2012). This allowed us to measure the photometry for each arc in about 16 bands from UV to near-infrared. For the remaining targets we used archival *HST* images; each source has imaging available in at least four bands except A1413, for which only two bands are available.

All of the selected targets have publicly available *Spitzer* IRAC observations, but due to the faintness of the gravitational arcs not all of them are detected. The arcs with useful IRAC imaging are A611, RXJ2129, A773, and MACS0717. We used channel 1 (3.6 μm) and channel 2 (4.5 μm) observations from *Spitzer* program 60034 (PI: E. Egami) for all of the arcs, and data from program 83 (PI: G. Rieke) for A773.

Additionally, some ground-based near-infrared data have been used for a few arcs: VLT ISAAC K_S-band for A1835 (Richard et al. 2006) and A1689 (J. Richard et al., in preparation), and Subaru MOIRCS H-band for A1703 (Richard et al. 2009).

2.3.2. *Palomar Observations*

Using ground and space-based archival data we can probe the spectral region redward of the Balmer break for each arc, with the exclusion of A1413. Since this type of photometry is crucial for measuring the stellar mass, we took K_S-band imaging for A1413 using WIRC on the Palomar 200-inch telescope. We used a 9-point dithering pattern of two-minute frames for a total exposure time of 198 minutes.

3. PROPERTIES OF STELLAR POPULATIONS

3.1. *Photometry*

Since the arcs experience a large gravitational magnification, they tend to lie at short projected distances from foreground galaxies. It may then be necessary to subtract the light of the cluster galaxies in order to reliably measure the arc photometry. In these cases we used Galfit (Peng et al. 2002), which allowed us to simultaneously fit many galaxies while taking into account the Point Spread Function (PSF) of the instrument, which was measured from bright, isolated stars in the field. This is particularly important when working with *Spitzer* IRAC images, which present a very large and asymmetric PSF.

Sometimes the gravitational arc emission is contaminated by the cluster bright central galaxy (BCG) luminous halo, and as a result the local background is hard to model. In these cases we fit the BCG using the Multi-Gaussian Expansion algorithm developed by Cappellari (2002). This method consists of fitting the surface brightness of a galaxy with a series of two-dimensional Gaussian functions, and is very effective for bright, extended galaxies.

After the subtraction of the foreground galaxies, the arc photometry is measured using polygonal apertures. The major source of uncertainty is generally the modeling of foreground galaxies and, for some IRAC images of crowded fields, confusion. The relative error in the flux is ~30% in the worst cases, but typically much less than that. All the photometric measurements are corrected for galactic extinction according to the map of Schlegel et al. (1998).

The photometric points are plotted for each arc in Figure 1 with a different color for each set of observations: red for *Hubble Space Telescope*, purple for ground-based near-infrared, and green for *Spitzer* data. When two gravitational images corresponding to the same source are available, the photometric analysis has been carried out independently on the different images, and the one less affected by foreground contamination was selected for the SED fitting. In these cases the name and the magnification factor of the chosen image are reported in Figure 1.

The WIRC K_S-band data for A1413 are not deep enough to detect the faint gravitational arc, because the contamination from the foreground BCG galaxy's halo is very strong. In this case we can only derive an upper limit on the flux, and we show it in Figure 1.

3.2. *SED Fitting*

For each target we fit the photometry from the observed UV to infrared using the stellar population models of Bruzual & Charlot (2003) in order to measure the stellar mass and other physical properties of the galaxies. We performed the fit using the chi-square minimization code FAST (Kriek et al. 2009) assuming the Calzetti et al. (2000) dust extinction law and a Chabrier (2003) initial mass function.

Since the emission lines detected in the Triplespec spectra are extremely bright compared to the continuum, we subtracted the measured line fluxes (see Section 4) from the appropriate photometric bands. This correction can be as high as 40% for a single band, but its effect on the stellar mass estimate is generally small because of the large number of photometric points used in the fit. For targets with multiple images on the slit the correction is calculated taking the flux from the combined spectrum, which has the advantage of a better signal to noise ratio, and then appropriately scaling it using the *HST* photometry.

Various stellar population parameters and degeneracies are involved in the process of SED fitting. One of the most important assumptions, and one that strongly affects the best-fit current SFR, is the star formation history. The widely used exponentially declining star formation history, or τ-model, may not be an appropriate choice since these galaxies are young and have a large gas reservoir. Sometimes an inverted τ-model is used for star-forming galaxies at high redshift. Both of these models require strong assumptions on the current state of the galaxy, i.e. that its star formation is currently at its minimum, or maximum, respectively. Also, they both introduce the free parameter τ which is usually not well-constrained by the data. For exponentially declining star formation histories, models with $\tau < 300$ Myr can give a formally acceptable fit to the data but usually fail in reproducing the SFR derived using other indicators (Wuyts et al. 2011). Since star-forming galaxies at high redshift are young, large values of τ imply a nearly flat star formation history. For these reasons we make the simplifying assumption of a constant star formation history. Shapley et al. (2005b) consider a sample of star forming galaxies at $z \sim 2$ and show that the agreement between stellar masses derived assuming τ-models or constant star formation history is very good, with no systematic offset and negligible dispersion. They also conclude that the choice of a particular SFH does not affect the uncertainty in the stellar mass measurement.

As shown by Wuyts et al. (2012a), the SED fit of low-mass star-forming galaxies tends generally to favor extremely young ages. Since a galaxy age cannot be smaller than the dynamical timescale, a lower limit on the SED age is often set. Following Wuyts et al. (2012a) we use 70 Myr as a lower limit, and the age of the Universe at the observed redshift as an upper limit. The effect of the age limit is not critical: lowering it to 20 Myr causes an average increase of 0.10 dex in both stellar mass and star formation rate.

One of the parameters involved in the SED fitting is the metallicity of the stellar population. This is different from the gas-phase metallicity, that we measure from rest-frame optical emission lines (see Section 4.3) and that can be higher than the stellar metallicity. The allowed values of stellar metallicity for the SED fitting are 0.2, 0.4 and 1.0 in units of solar metallicity. For some of the arcs, the best-fit value is significantly larger than the gas-phase metallicity. We attribute this unphysical result to the effect of SED fitting degeneracies. We performed a test for each arc by fitting the SED while keeping the metallicity fixed at the value closest to the one measured via emission lines. The resulting stellar population parameters are negligibly affected, and their uncertainties change only marginally.

It is important to note that among the SED fitting output parameters, stellar mass is the most robust (e.g. Wuyts et al. 2007), and is also the only one that is critical for our analysis. Star formation rate and dust extinction are more sensitive to the assumptions made, but it is possible to compare them with independent measurements from the rest-frame optical emission lines.

The best-fit spectra are shown in Figure 1, and the output parameters (stellar mass, dust extinction, age, and current star formation rate) for each arc are listed in Table 4. Stellar masses and star formation rates are corrected for the gravitational magnification. The stellar masses are in the range $7.8 < \log M/M_\odot < 9.4$, and are located at the low end of the mass distribution of the SDSS sample (e.g. Zahid et al. 2012a). The uncertainties are between 0.1 and 0.3 dex except for A1413, for which the low number of photometric points yields a large uncertainty in the stellar mass, $\Delta \log M = 0.36$ dex.

It is common practice to report the best-fit stellar mass (i.e. the one corresponding to the model that best describes the photometric data) and the 68% confidence region. The error bars are often highly asymmetric, and are very difficult to propagate when using the SED results in further analysis. In fact, a rigorous propagation of asymmetric error bars is possible only when the posterior distribution is known. Instead, we calculate the stellar mass posterior distribution from the chi-square grid output from FAST, and report the mean and the standard deviation of the distribution. The posterior distributions in $\log M$ are only weakly skewed and are well approximated by a Gaussian function. On the other hand, the best-fit value is often off-center, and choosing it as the best estimate would cause asymmetric error bars. Reporting the mean and standard deviation has the advantage of a straightforward propagation of the uncertainty in following calculations, which is essential for the present work. The same arguments apply to other stellar population parameters such as $\log$ SFR and dust extinction $E(B-V)$, and we follow the same method for estimating their values. The stellar population age, however, presents a posterior distribution that is very skewed for those galaxies with a best-fit age near the lower limit, therefore we list the best-fit value and the asymmetric 68% confidence interval, which we do not use in any further analysis.

4. SPECTROSCOPIC DIAGNOSTICS

The goal of this study is to explore the relation between stellar mass, star formation rate and gas metallicity for star-forming galaxies at high redshift. In the previous section we derived the stellar masses using photometric data. In this section we use the rest-frame optical emission lines of the gravitational arcs to measure their star formation rate and metallicity.

TABLE 3
Emission line fluxes

Source	[OII]$\lambda3727, 3729$	[NeIII]$\lambda3870$	Hβ	[OIII]$\lambda4959$	[OIII]$\lambda5007$	Hα	[NII]$\lambda6584$
A611	$\cdots$	$\cdots$	15.9 ± 4.5	$34.0 \pm 8.6^*$	90.4 ± 4.1	43.3 ± 6.0	$< 4.1^*$
RXJ2129	$\cdots$	$\cdots$	$10.6 \pm 5.1^*$	$11.8 \pm 4.0^*$	45.8 ± 7.3	32.8 ± 5.9	< 3.7
A1413	$< 29^*$	10.5 ± 3.7	$12.8 \pm 3.2^*$	$60.8 \pm 6.3^*$	176 ± 3	97.2 ± 3.7	$< 11^*$
A1835	36 ± 13	5.8 ± 2.1	16.6 ± 3.1	$< 14^*$	44.3 ± 4.2	45.4 ± 8.1	5.5 ± 2.1
A773	23.8 ± 9.7	< 24	$8.3 \pm 3.5^*$	15.3 ± 3.4	49.0 ± 5.2	47.2 ± 7.6	< 9.2
MACS0717	$19.3 \pm 9.0^*$	< 9.1	8.5 ± 2.9	4.4 ± 1.7	18.3 ± 2.3	18.9 ± 4.9	< 6.0
A383	18.5 ± 5.9	$< 38^*$	$< 8.5^*$	$< 6.9^*$	16.4 ± 2.2	25.2 ± 7.1	< 7.4
A1689	$< 29^*$	$< 7.7^*$	$< 11^*$	$< 24^*$	49.1 ± 4.9	$\cdots$	$\cdots$
A1703	12.7 ± 5.7	5.5 ± 1.0	11.2 ± 2.5	34.0 ± 3.6	81.1 ± 2.0	$\cdots$	$\cdots$

NOTE. — Fluxes in units of 10^{-17} erg s^{-1} cm^{-2}. The listed uncertainties apply to flux ratios, and do not include the error in the absolute flux calibration. For undetected lines the 2-σ upper limit is given.
* Lines strongly contaminated by sky emission lines. These measurements are to be considered highly uncertain.

In order to derive physical quantities from the observed spectra, we need to quantitatively analyze the emission lines. Each emission line profile was fitted with one Gaussian (two for the doublet [OII]$\lambda3727, 3729$). For each line we derived flux, redshift, width and continuum level from the fit. For the faintest lines we fixed one or more of these parameters using as a reference [OIII]$\lambda5007$ or Hα, which have a relatively high signal to noise ratio. The line fluxes are reported in Table 3. The redshifts, measured from [OIII]$\lambda5007$, are given in Table 1. In two cases we found that previously published redshifts were incorrect (A611 and RXJ2129, Richard et al. 2010) due to misidentification of rest-frame UV spectral features.

4.1. Dust Extinction

We estimated dust reddening using the Hα/Hβ flux ratio. Assuming a case B recombination and typical temperature (10,000 K) and density (100 cm^{-3}), the theoretical value of the ratio is 2.86 (Osterbrock & Ferland 2006). We used the Calzetti et al. (2000) law to derive the dust extinction from the observed flux ratio. The results are shown in Figure 2 and compared to the dust extinction derived from the SED fitting. The two methods are in good agreement and provide very low dust extinction for most of the arcs. Since the Balmer decrement is highly sensitive to the Hβ flux, we plot with a different symbol the arcs for which this line is contaminated by OH sky lines. The only object with a large discrepancy between the two measurements of dust extinction is A1413, for which the Balmer decrement gives $E(B-V) \sim 0.8$. We attribute this very large value to the effect of sky emission on the Hβ flux, since all the other galaxies in our sample have $E(B-V) < 0.3$. The SED fit for this galaxy, although uncertain, being based on only two photometric points, gives a dust extinction very similar to that found for the other arcs, $\langle E(B-V) \rangle = 0.13$.

The Balmer decrement probes the extinction in the HII regions, where the nebular emission originates, while the SED fit output applies to the overall stellar population of a galaxy. In principle by comparing the dust extinction obtained by the two methods it is possible to study the dust distribution, which can be concentrated in star-forming regions. Calzetti et al. (1994) found that for local starburst galaxies the nebular dust extinction is roughly twice the extinction of the stellar continuum. At $z \sim 2$ it is not clear whether there is a difference between the reddening experienced by stellar and gas emission (e.g. Erb et al. 2006b; Hainline et al. 2009;
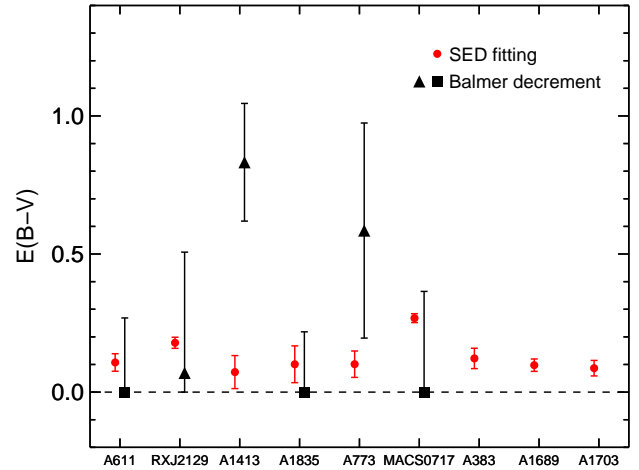

FIG. 2.— Dust extinction E(B-V) derived from SED fitting (red circles), Balmer decrement (black squares), and Balmer decrement with Hβ affected by sky line contamination (black triangles).

Förster Schreiber et al. 2009). Our data suggest a similar amount of attenuation for the two components, but the Hβ flux determinations are too noisy to draw any conclusion. When in the following analysis we correct the emission line fluxes for dust extinction, we always use the SED fitting values, which are less affected by uncertainty. This method could underestimate the dust extinction experienced by gas emission by a factor of 2, which translates into an average increase in SFR of only 0.15 dex.

4.2. Lack of AGN Contribution

To exclude the possibility of any AGN contribution to the gravitational arc emission, in Figure 3 we show the [OIII]$\lambda5007$/Hβ vs. [NII]$\lambda6584$/Hα line ratio diagram (BPT diagram, Baldwin et al. 1981). In this plot star-forming galaxies and AGN populate separate regions due to the different ionization mechanisms at the origin of the line emission. All the gravitational arcs lie on or near the star-forming branch of the diagram, and we can firmly exclude the presence of AGN in our sample.

It is interesting to note that the location of these $z \sim 2$ star-forming galaxies on the BPT diagram is not coincident with the region most populated by low-redshift galaxies. In particular, none of our objects lie in the region $\log([\text{OIII}]\lambda5007/\text{H}\beta) < 0$, where the majority of local galaxies are found. A large [OIII]$\lambda5007$/ Hβ ratio has already been reported in many studies of high-redshift

TABLE 4
Physical properties of the sample

Source	$\log(M/M_\odot)$[a]	E(B-V)[a]	$\log(\text{Age/yr})$[a]	SFR_{SED}[a] $(M_\odot/\text{yr})$	$\text{SFR}_{\text{H}\alpha}$[b] $(M_\odot/\text{yr})$	Line width[b] (km/s)	$12 + \log(\text{O/H})$[b]	FMR residual[c]
A611	8.27 ± 0.09	0.11 ± 0.03	$8.3^{+0.3}_{-0.4}$	1.7 ± 0.5	2.0 ± 1.1	27 ± 5	7.89 ± 0.19	-0.10 ± 0.21
RXJ2129	7.80 ± 0.25	0.18 ± 0.02	$7.9^{+0.1}_{-0.0}$	0.9 ± 0.5	0.6 ± 0.4	39 ± 10	7.89 ± 0.40	0.05 ± 0.43
A1413	8.72 ± 0.36	0.07 ± 0.06	$8.0^{+0.5}_{-0.1}$	3.1 ± 1.7	7.2 ± 4.3	30 ± 3	7.89 ± 0.33	-0.24 ± 0.38
A1835	8.33 ± 0.22	0.10 ± 0.07	$8.8^{+0.4}_{-1.0}$	1.1 ± 0.7	1.0 ± 0.7	54 ± 7	8.45 ± 0.07	0.38 ± 0.15
A773	9.16 ± 0.21	0.10 ± 0.05	$8.7^{+0.4}_{-0.5}$	4.6 ± 2.7	4.4 ± 2.7	50 ± 7	8.32 ± 0.11	-0.08 ± 0.17
MACS0717	9.36 ± 0.19	0.27 ± 0.02	$7.9^{+0.1}_{-0.1}$	34 ± 15	15 ± 10	65 ± 8	8.53 ± 0.10	0.11 ± 0.16
A383	8.67 ± 0.16	0.12 ± 0.04	$8.1^{+0.6}_{-0.3}$	3.4 ± 1.3	3.8 ± 2.3	54 ± 8	8.56 ± 0.10	0.41 ± 0.14
A1689	8.27 ± 0.24	0.10 ± 0.02	$8.0^{+0.3}_{-0.2}$	2.2 ± 1.3	$\cdots$	80 ± 9	7.89 ± 0.39	-0.10 ± 0.41
A1703	8.49 ± 0.27	0.09 ± 0.03	$8.6^{+0.7}_{-0.7}$	1.0 ± 0.5	4.1 ± 2.9	34 ± 3	7.84 ± 0.14	-0.22 ± 0.22

Note. — Stellar masses and star formation rates are corrected for the gravitational magnification.
[a] Derived from SED fitting.
[b] Derived from rest-frame optical emission lines.
[c] Difference between the measured metallicity and the metallicity predicted by the FMR as formulated by Mannucci et al. (2011).

star-forming galaxies (Shapley et al. 2005a; Erb et al. 2006a, 2010; Hainline et al. 2009; Richard et al. 2011; Rigby et al. 2011), and is indicative of a high ionization parameter, as extensively discussed by Erb et al. (2010).

4.3. Metallicities

Rest-frame optical nebular lines contain a large amount of information on the physical conditions of the gas responsible for the emission, including its metallicity. If the auroral line [OIII]λ4363 is detected, then it is possible to calculate the electron temperature and have a direct measurement of the metallicity. Unfortunately this line is so weak that at high redshift it has been detected only for a handful of objects. Instead, we derived the gas metallicity from the flux ratio of strong emission lines. There are several well-established methods to estimate the metallicity from flux ratios, calibrated using either theoretical calculations or observations of low-redshift galaxies. The absolute metallicity obtained with these methods is highly uncertain, and the different sets of calibrations,
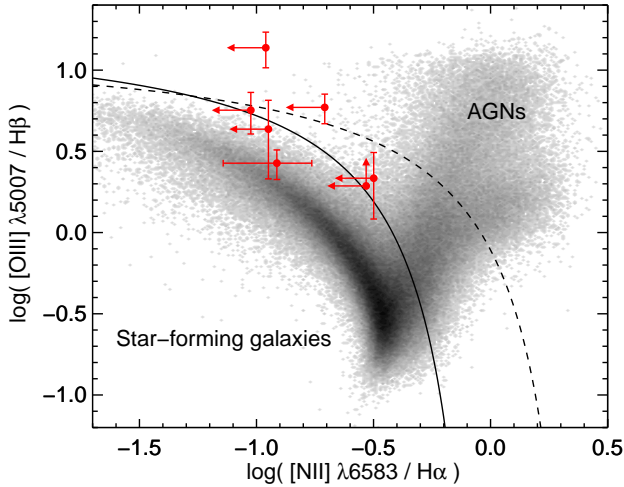

Fig. 3.— BPT diagram (Baldwin et al. 1981). Our sample (red points) is compared to the SDSS sample (gray density map, Kauffmann et al. 2003). Also shown are the theoretical (dashed line, Kewley et al. 2001) and empirical (solid line, Kauffmann et al. 2003) separation between active galactic nuclei and star-forming galaxies.

can differ by as much as 0.7 dex (Kewley & Ellison 2008). Although this discrepancy makes it very difficult to compare observational results obtained with different calibrations, relative measurements obtained with the same strong line method are much more reliable.

The main goal of the present study is to test whether the locally determined fundamental metallicity relation applies to high-redshift galaxies as well. The natural choice is then to use the same metallicity calibrations adopted by Mannucci et al. (2010) in the definition of the local relation. These are the empirical calibrations of Maiolino et al. (2008), which give a polynomial fit for the value of various nebular line ratios as a function of the gas metallicity. The main line ratios are [OIII]λ5007/Hβ and [OIII]λ5007/[OII]λ3727, 3729, while [NII]λ6584/Hα and [NeIII]λ3870/[OII]λ3727, 3729 were used only for some arcs, mostly as upper or lower limits. All the arc metallicities were calculated using at least two line ratios, which correspond to a minimum of three independent line fluxes. This minimizes the effects of catastrophic errors due to sky line contamination. When many line ratios are available for the same object, they are always consistent with each other, within the observational uncertainties.

It is worth remarking that the ratio between the flux of two lines is independent of the absolute flux calibration even for lines that lie in very distant parts of the spectrum, since Triplespec allows us to observe the J, H and K band simultaneously. It is also independent of gravitational magnification and slit loss. We corrected the line fluxes for dust extinction, using the SED fitting results, in a differential way: the ratio of two lines depends only on the ratio of the attenuation at the corresponding wavelengths. This results in a small correction to the metallicity estimate and its uncertainty. The dust-corrected flux for Hα was also used as a proxy for Hβ (assuming Hα/Hβ = 2.86) in order to have an additional constraint.

The relation between [OIII]λ5007/Hβ and the metallicity is not monotonic, and presents a maximum at $12 + \log(\text{O/H}) = 7.89$. Since this is a stationary point, any uncertainty in the line ratio is transformed into a much larger uncertainty in the metallicity. About half of our sample is found in this location, with large un-
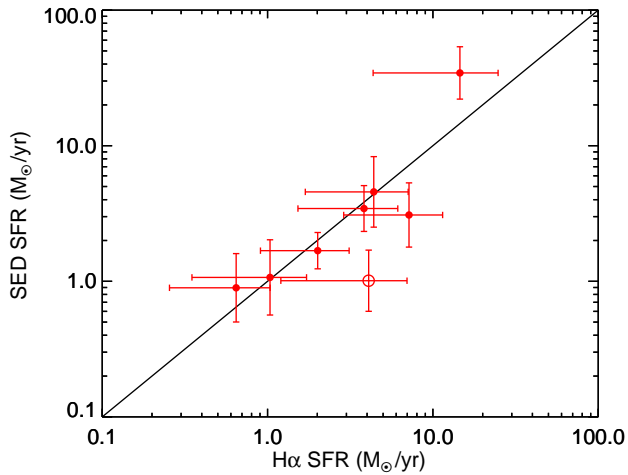
FIG. 4.— Comparison of star formation rates derived using two methods: SED-fitting and Hα flux. The empty circle is A1703 for which Hβ has been used as a proxy for Hα.

certainties on log(O/H), up to 0.4 dex. The shape of this calibration also causes the existence of two possible metallicity values in some cases, but an unambiguous solution is always found thanks to the other line ratios.

The final gas metallicity for each arc is taken to be the weighted average of the results from all the available line ratios and is given in Table 4.

The metallicity of the two gravitational arcs A1689 and A1835 has already been measured by Richard et al. (2011) from near-infrared spectra obtained with Keck NIRSPEC and using the same set of Maiolino et al. (2008) calibrations. Our results are in good agreement.

4.4. *Star Formation Rates*

We derived the current star formation rate from the extinction-corrected Hα emission flux using the calibration given by Kennicutt (1998), dividing the result by 1.7 to convert to that appropriate for a Chabrier IMF. The resulting star formation rates, corrected for the gravitational magnification, are reported in Table 4.

The SFR derived from nebular emission accounts for the star formation activity in the physical region of the arc that is covered by the slit, which is different, in principle, from the star formation rate of the entire galaxy. But the narrow gravitational arcs from our sample are easily covered by the 1 arcsec wide slit, as is apparent from the image stamps in Figure 1. Therefore we do not attempt to correct for this effect, which is in any case less important than the uncertainty caused by slit alignment and seeing variability.

Figure 4 shows excellent agreement between the star formation rates calculated using SED fitting and Hα flux. This is encouraging because it validates the numerous assumptions made in the derivation of the star formation rates. It is particularly interesting that the agreement between SED-fitting and nebular emission even holds for A1413, where only two photometric points are available. It is possible that the simplifying choice of a constant star formation history, with the consequent decrease in the number of free parameters, helped reduce the scatter in the comparison between the two methods.

In the following section we will always use the star formation rate derived from the Hα flux. The spectra of the two objects at $z > 3$ do not include Hα, which is

redshifted outside the Triplespec range. For one of them (A1703, empty circle in Figure 4) we use the observed flux of Hβ as a proxy for Hα, assuming the theoretical line ratio discussed in Section 4.1 and correcting for dust extinction. For A1689, for which both Hα and Hβ are not available, we use the SED fitting star formation rate.

4.5. *Line Widths*

The broadening of the emission lines due to the gas kinematics depends on the gravitational well of the galaxies. Measuring the line widths can then give an estimate on the gravitational arc masses that is independent of SED fitting and gravitational lensing models.

The observed velocity width of the nebular lines need to be corrected for the instrumental resolution, which was measured from the sky OH lines. For each source we extracted the spectrum of the sky using the same procedure followed to extract the arc spectrum. We measured the dispersion of the brightest, unblended OH lines, which is 40-55 km s^{-1} depending on spectral order and wavelength. We calculated a linear fit of the ratio of the spectral resolution R to the order m as a function of wavelength, and used this to estimate the instrumental resolution for each nebular line.

Most of the arc emission lines are well-resolved. For each source we take the weighted mean of the line widths of all the well-detected lines excluding [OII]$\lambda 3727, 3729$, which is a doublet and is not completely resolved. The results are listed in Table 4.

Since a detailed lensing map is needed to measure the intrinsic radius of the gravitational arcs, we do not attempt to estimate the dynamical masses. The observed velocity dispersions, however, are unusually low if compared to the results of similar studies (Law et al. 2009; Förster Schreiber et al. 2009; Jones et al. 2010). This is an important confirmation of the low masses found in our sample.

5. RESULTS

In this section we combine our measurements of stellar mass, gas metallicity and star formation rate to explore the properties of our sample of low-mass galaxies. In particular, the goal of this work is to test whether high-redshift galaxies follow the local fundamental metallicity relation, claimed to be valid up to at least $z \sim 2.2$ (Mannucci et al. 2010).

5.1. *The Mass–Metallicity Relation*

The mass-metallicity relation for our sample is plotted in Figure 5 together with the fit to the local relation from Kewley & Ellison (2008) and to the high-redshift one of Erb et al. (2006a). All the results shown in this figure have been derived using the same set of metallicity calibrations (the fits shown are taken from Maiolino et al. 2008, and are corrected for the choice of IMF).

The magnification caused by gravitational lensing allows us to probe stellar masses much smaller than those considered in previous studies of unlensed galaxies, even in the nearby Universe. This makes a direct comparison difficult, but it is clear from Figure 5 that our points do not lie on the extrapolation of neither the $z \sim 0$ nor the $z \sim 2$ relations, and have a substantial scatter, larger than the observational uncertainties.
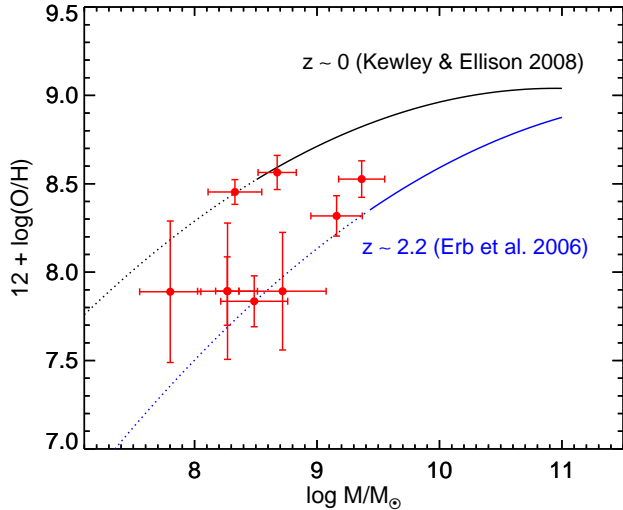
FIG. 5.— Mass-metallicity relation for our sample (red points). The fit to the relation at $z \sim 0$ (black solid line, Kewley & Ellison 2008) and $z \sim 2.2$ (blue solid line, Erb et al. 2006a) are also shown. Note that these fits are calculated using the Maiolino et al. (2008) metallicity calibrations. The dashed lines are extrapolation at low masses.
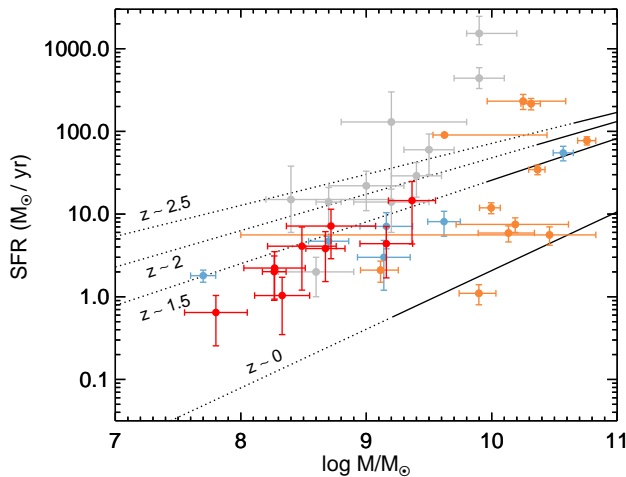

FIG. 6.— Star formation rate versus stellar mass diagram. The main sequence of star-forming galaxies is shown in black at $z \sim 0$ (Zahid et al. 2012b) and at high redshift (Whitaker et al. 2012), with dotted lines being the extrapolation at masses below the completeness limit. Our sample (red points) is roughly on the main sequence at $z \sim 2$. Points from other studies of lensed galaxies at high redshift are shown: Richard et al. (2011) in orange, Wuyts et al. (2012b) in gray, and Christensen et al. (2012) in blue.

Yuan et al. (2013) measured the mass-metallicity relation for a sample of gravitational arcs at $z \sim 2$ obtained by combining their data with results from previous studies. Although a comparison of the absolute measurements is not possible because they use a different metallicity calibration, their results are similar to ours: the lensed galaxies tend to have lower metallicities than the SDSS galaxies but do not lie on a tight sequence.

This results are consistent with the hypothesis of a mass-metallicity relation dependent on some other parameter that is not necessarily the redshift. In the remaining parts of this section we will explore the role of the star formation rate.

5.2. Star Formation Rate versus Stellar Mass

In Figure 6 we show the location of our sample in the SFR-stellar mass diagram (red points), compared to the results of other studies of lensed galaxies at high redshift: Richard et al. (2011, in orange), Wuyts et al. (2012b, in gray), and Christensen et al. (2012, in blue). Our sample populates the lower left corner, with masses and star formation rates on average lower than what probed by previous studies. In particular, we more than doubled the number of low-mass galaxies ($M < 10^9 M_\odot$) at this redshift with known metallicity and SFR.

In the mass-SFR plane, star-forming galaxies lie on a relatively tight relation often called the main sequence (Noeske et al. 2007). This sequence evolves strongly with redshift, with the normalization decreasing over cosmic time at least since $z \sim 2.5$ (Whitaker et al. 2012; Zahid et al. 2012b, black lines in Figure 6). The gravitational arcs that we selected have star formation rates that are on or below the main sequence at $z \sim 2$. Previous studies of lensed galaxies did not reach such low values of SFR, with the exception of the work of Christensen et al. (2012).

The very low star formation rate of these arcs, between 0.6 and 15 $M_\odot$/yr, is of fundamental importance in this study. First, the fact that our sample lies on the main sequence means that these galaxies are representative of the typical population of star-forming galaxies. Shallower studies are biased towards luminous galaxies with star formation rates much higher than the main sequence. These objects are thought to be in a starburst phase, potentially caused by a merger, and are not representative of the typical conditions of star-forming galaxies. Secondly, it allows us to compare high and low redshift galaxies with similar star formation rates, and thereby directly address the goals of this paper. From Figure 6 we can see that local massive galaxies ($9.5 < \log M/M_\odot < 11$) that lie on the main sequence at $z \sim 0$ have SFRs comparable to our sample. This is not the case for the majority of lensed galaxies considered by previous studies.

5.3. The Fundamental Metallicity Relation

We now turn our attention to the fundamental metallicity relation, a surface in the 3D parameter space of stellar mass, metallicity and star formation rate tightly followed by the SDSS galaxies (Mannucci et al. 2010). In Figure 7 we plot for each gravitational arc the difference between the metallicity that we measure from nebular lines and the metallicity predicted by the local FMR given its stellar mass and star formation rate. We also show the points from the previous studies of lensed galaxies. Clearly there is a correlation between metallicity residuals and stellar mass: at low masses the FMR underestimates the metallicity. The mean metallicity residual for our sample is 0.37 ± 0.09 dex.

At such low masses, even the SDSS sample has few objects. In fact we are using here an extrapolation of the fundamental metallicity relation, which was originally found for galaxies more massive than $10^{9.2} M_\odot$. Mannucci et al. (2011) extended the local FMR down to $10^{8.3} M_\odot$, finding a slightly different slope and a larger scatter. With this version of the fundamental metallicity relation the residuals, shown in Figure 8, have a much smaller dispersion. The weighted average of the low-mass FMR residuals is 0.12 ± 0.07. However, the
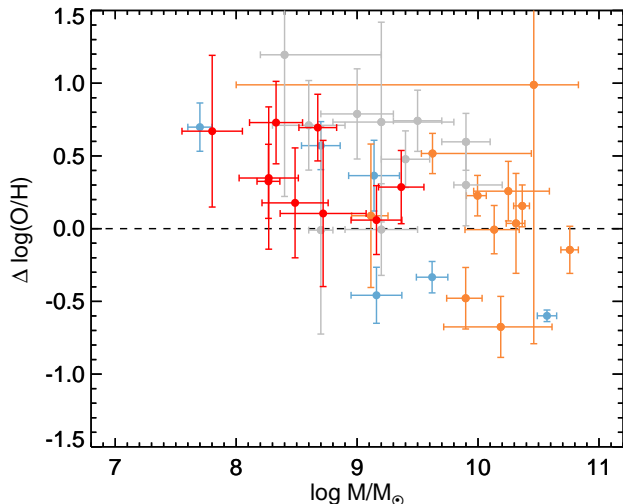
FIG. 7.— Difference between measured metallicity and the prediction of the fundamental metallicity relation from Mannucci et al. (2010). The points are color-coded as in Figure 6.
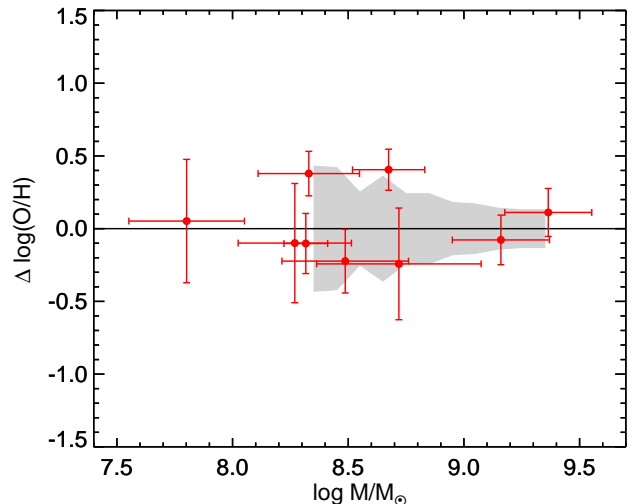

FIG. 8.— Difference between measured metallicity and the prediction of the FMR extension to low masses (Mannucci et al. 2011). Our points are shown in red (with A611 offset by 0.05 dex in mass for clarity). The standard deviation of the SDSS low-mass galaxies is shown in gray.

weighted mean is skewed toward galaxies with higher metallicity since they are measured with higher precision, because the metallicity calibrations are not linear. This is clear from Figure 5, where the group of five galaxies aligned at $12 + \log(\mathrm{O/H}) \sim 7.9$, the maximum of the [OIII]$\lambda 5007$/Hβ line ratio calibration, present the largest error bars. The arithmetic mean is not affected by this bias, and gives $\langle \Delta \log(\mathrm{O/H}) \rangle = 0.02 \pm 0.08$. The agreement of high-redshift lensed galaxies with the local fundamental metallicity relation is remarkable, and strongly suggests that these objects lie on the relation independently of their redshift at least up to $z \sim 3$. Previous studies of lensed galaxies also found a general agreement, but without conclusive results for $M < 10^9 M_{\odot}$, perhaps because of a less strict selection in SFR. Also, this is the first time that the universality of the FMR is confirmed using galaxies at high redshift with a SFR which is observed in typical galaxies in the local Universe.

From Figure 8 it is clear that our gravitational arcs show a relatively small scatter around the local fundamental metallicity relation. The standard deviation of the metallicity offsets from the FMR is 0.24 dex, and the mean error in $\Delta \log(\mathrm{O/H})$ is 0.25 dex. Therefore the observed scatter could be in principle just a product of observational uncertainties. In contrast, the standard deviation found by Mannucci et al. (2011) for the SDSS sample is about 0.4 dex at $10^{8.4} M_{\odot}$, and is shown in gray in Figure 8 as a function of stellar mass. Although they do not report the typical errors on mass, star formation rate and metallicity, they claim that the observational uncertainties are not large enough to explain the observed dispersion. Furthermore, SDSS galaxies are selected by requiring a signal to noise ratio greater than 25 for the Hα flux, therefore the uncertainty in their metallicity must be much smaller than for our sample. Since the uncertainty in $\Delta \log(\mathrm{O/H})$ is mainly due to the metallicity measurement, we conclude that the intrinsic scatter at $z \sim 2$ is most likely smaller than at $z = 0$.

6. DISCUSSION

Our data confirm that the fundamental metallicity relation applies to low-mass galaxies at $1.5 < z < 3$. This suggests that this relation is time-invariant and therefore universal.

In the equilibrium model of Finlator & Davé (2008), metallicity and star formation rates are tightly connected to gas inflows and outflows, so that a change in one implies a consequent change in the other (see also Davé et al. 2011a, 2012). If each of these processes are in equilibrium, then the FMR is naturally explained. If a galaxy is perturbed, by e.g. a minor merger, after a certain time it will return to the equilibrium configuration. The observed evolution in the mass-metallicity relation could be due to the fact that we are sampling galaxy populations with different star formation rates at different redshifts. This would explain why our points do not lie on the low-mass end of the $z \sim 2$ mass-metallicity relation from Erb et al. (2006a), that was determined using relatively high-SFR galaxies.

The low scatter that we found at high redshift may provide a valuable additional constraint for numerical or analytical models of galaxy evolution. This is particularly important because at low masses models have diverging predictions (Zahid et al. 2012a). In the equilibrium model, the observed scatter of the FMR is determined by how quickly a perturbed galaxy can return to equilibrium. This timescale, in turn, depends on the mass loading factor, a parameter that is fundamental for hydrodynamic simulations. The observation of the scatter in the FMR at different stellar masses and redshifts therefore gives important constraints on numerical simulations of galaxy evolution. But current simulations do not resolve stellar masses below $10^9 M_{\odot}$ (e.g. Davé et al. 2011b).

An interesting perspective on the issue of the FMR scatter has been pointed out by Zahid et al. (2012a). They studied the mass-metallicity relation at $z \sim 0$ using various samples from the literature and found a clear increase in the intrinsic scatter at low stellar masses. They suggest the possibility that this scatter is due to

a population of low-mass, metal-rich galaxies which are near the end of their star formation. At a fixed mass, the same amount of metals would give a higher metallicity measurement, since there is little gas left. Assuming that the scatter in the FMR is directly caused by the scatter in the mass-metallicity relation, our observations agree well with this scenario, since at high redshift such a population of low-mass galaxies terminating their star-formation would not be expected.

7. SUMMARY

We present near-infrared spectroscopic data for 9 gravitational arcs between redshift 1.5 and 3.3, and the measurement of their stellar mass, gas metallicity and star formation rate. The use of strong gravitational lensing allows us to probe very low masses and star formation rates. Our sample more than doubles the number of galaxies with stellar masses below $10^9 M_\odot$ at $z \sim 2$ with known metallicity and SFR. Our main goal is to test whether these galaxies follow the fundamental metallicity relation discovered for local galaxies.

We find that the gravitational arcs lie above the mass-metallicity relation at $z \sim 2$ but below the local relation. However, they also have SFRs that are roughly on the main sequence of star-forming galaxies. This means that they are representative of typical star-forming galaxies, i.e. they are not in a starburst phase.

Our data are fully consistent with the low mass extension of the local fundamental metallicity relation (Mannucci et al. 2011), with a mean metallicity offset of 0.02 ± 0.08 dex. The dispersion around the FMR of 0.24 dex is smaller than the one measured for local galaxies, and represents an important additional constraint for galaxy evolution models.

REFERENCES

Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, ApJ, 533, 682
Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582
Cappellari, M. 2002, MNRAS, 333, 400
Chabrier, G. 2003, PASP, 115, 763
Christensen, L., et al. 2012, MNRAS, 427, 1953
Coe, D., Benítez, N., Broadhurst, T., & Moustakas, L. A. 2010, ApJ, 723, 1678
Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
Cushing, M. C., Vacca, W. D., & Rayner, J. T. 2004, PASP, 116, 362
Davé, R., Finlator, K., & Oppenheimer, B. D. 2011a, MNRAS, 416, 1354
—. 2012, MNRAS, 421, 98
Davé, R., & Oppenheimer, B. D. 2007, MNRAS, 374, 427
Davé, R., Oppenheimer, B. D., & Finlator, K. 2011b, MNRAS, 415, 11
Dayal, P., Ferrara, A., & Dunlop, J. S. 2012, ArXiv e-prints
De Lucia, G., Kauffmann, G., & White, S. D. M. 2004, MNRAS, 349, 1101
Elias, J. H., Frogel, J. A., Matthews, K., & Neugebauer, G. 1982, AJ, 87, 1029
Ellison, S. L., Patton, D. R., Simard, L., & McConnachie, A. W. 2008, ApJ, 672, L107
Erb, D. K., Pettini, M., Shapley, A. E., Steidel, C. C., Law, D. R., & Reddy, N. A. 2010, ApJ, 719, 1168
Erb, D. K., Shapley, A. E., Pettini, M., Steidel, C. C., Reddy, N. A., & Adelberger, K. L. 2006a, ApJ, 644, 813
Erb, D. K., Steidel, C. C., Shapley, A. E., Pettini, M., Reddy, N. A., & Adelberger, K. L. 2006b, ApJ, 647, 128
Finlator, K., & Davé, R. 2008, MNRAS, 385, 2181
Förster Schreiber, N. M., et al. 2009, ApJ, 706, 1364
Garnett, D. R. 2002, ApJ, 581, 1019
Hainline, K. N., Shapley, A. E., Kornei, K. A., Pettini, M., Buckley-Geer, E., Allam, S. S., & Tucker, D. L. 2009, ApJ, 701, 52
Herter, T. L., et al. 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7014, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
Jones, T. A., Swinbank, A. M., Ellis, R. S., Richard, J., & Stark, D. P. 2010, MNRAS, 404, 1247
Kauffmann, G., et al. 2003, MNRAS, 346, 1055
Kennicutt, Jr., R. C. 1998, ApJ, 498, 541
Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121
Kewley, L. J., & Ellison, S. L. 2008, ApJ, 681, 1183
Kriek, M., van Dokkum, P. G., Labbé, I., Franx, M., Illingworth, G. D., Marchesini, D., & Quadri, R. F. 2009, ApJ, 700, 221
Lara-López, M. A., et al. 2010, A&A, 521, L53
Larson, R. B. 1974, MNRAS, 169, 229
Law, D. R., Steidel, C. C., Erb, D. K., Larkin, J. E., Pettini, M., Shapley, A. E., & Wright, S. A. 2009, ApJ, 697, 2057
Lequeux, J., Peimbert, M., Rayo, J. F., Serrano, A., & Torres-Peimbert, S. 1979, A&A, 80, 155
Limousin, M., et al. 2011, ArXiv e-prints
Maiolino, R., et al. 2008, A&A, 488, 463
Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, MNRAS, 408, 2115
Mannucci, F., Salvaterra, R., & Campisi, M. A. 2011, MNRAS, 414, 1263
Mannucci, F., et al. 2009, MNRAS, 398, 1915
Newman, A. B., Treu, T., Ellis, R. S., & Sand, D. J. 2011, ApJ, 728, L39
Newman, A. B., Treu, T., Ellis, R. S., Sand, D. J., Richard, J., Marshall, P. J., Capak, P., & Miyazaki, S. 2009, ApJ, 706, 1078
Noeske, K. G., et al. 2007, ApJ, 660, L43
Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Postman, M., et al. 2012, ApJS, 199, 25
Richard, J., Jones, T., Ellis, R., Stark, D. P., Livermore, R., & Swinbank, M. 2011, MNRAS, 413, 643
Richard, J., Pei, L., Limousin, M., Jullo, E., & Kneib, J. P. 2009, A&A, 498, 37
Richard, J., Pelló, R., Schaerer, D., Le Borgne, J.-F., & Kneib, J.-P. 2006, A&A, 456, 861
Richard, J., et al. 2010, MNRAS, 404, 325
Rigby, J. R., Wuyts, E., Gladders, M. D., Sharon, K., & Becker, G. D. 2011, ApJ, 732, 59
Savaglio, S., et al. 2005, ApJ, 635, 260
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Shapley, A. E., Coil, A. L., Ma, C.-P., & Bundy, K. 2005a, ApJ, 635, 1006
Shapley, A. E., Steidel, C. C., Erb, D. K., Reddy, N. A., Adelberger, K. L., Pettini, M., Barmby, P., & Huang, J. 2005b, ApJ, 626, 698
Tassis, K., Kravtsov, A. V., & Gnedin, N. Y. 2008, ApJ, 672, 888

Tremonti, C. A., et al. 2004, ApJ, 613, 898

Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2004, PASP, 116, 352

Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJ, 754, L29

Wuyts, E., Rigby, J. R., Gladders, M. D., Gilbank, D. G., Sharon, K., Gralla, M. B., & Bayliss, M. B. 2012a, ApJ, 745, 86

Wuyts, E., Rigby, J. R., Sharon, K., & Gladders, M. D. 2012b, ApJ, 755, 73

Wuyts, S., et al. 2007, ApJ, 655, 51

—. 2011, ApJ, 738, 106

Yates, R. M., Kauffmann, G., & Guo, Q. 2012, MNRAS, 422, 215

Yuan, T.-T., Kewley, L. J., & Richard, J. 2013, ApJ, 763, 9

Zahid, H. J., Bresolin, F., Kewley, L. J., Coil, A. L., & Davé, R. 2012a, ApJ, 750, 120

Zahid, H. J., Dima, G. I., Kewley, L. J., Erb, D. K., & Davé, R. 2012b, ApJ, 757, 54

Zahid, H. J., Kewley, L. J., & Bresolin, F. 2011, ApJ, 730, 137